



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/071,934	02/07/2002	Markus Walther	9432-000151	9849

27572 7590 08/07/2006

HARNESS, DICKEY & PIERCE, P.L.C.
P.O. BOX 828
BLOOMFIELD HILLS, MI 48303

EXAMINER

NG, EUNICE

ART UNIT	PAPER NUMBER
----------	--------------

2626

DATE MAILED: 08/07/2006

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary

Application No.

10/071,934

Applicant(s)

WALTHER, MARKUS

Examiner

Eunice Ng

Art Unit

2626

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) ☒ Responsive to communication(s) filed on 23 June 2006.
- 2a) ☐ This action is **FINAL**. 2b) ☒ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) ☒ Claim(s) 1-19 is/are pending in the application.
- 4a) Of the above claim(s) _____ is/are withdrawn from consideration.
- 5) ☐ Claim(s) _____ is/are allowed.
- 6) ☒ Claim(s) 1-19 is/are rejected.
- 7) ☐ Claim(s) _____ is/are objected to.
- 8) ☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☐ The drawing(s) filed on _____ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.
- Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
- Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All b) ☐ Some * c) ☐ None of:
1. ☐ Certified copies of the priority documents have been received.
2. ☐ Certified copies of the priority documents have been received in Application No. _____.
3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).
- * See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

- 1) ☒ Notice of References Cited (PTO-892)
- 2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
- 3) ☐ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)
Paper No(s)/Mail Date _____
- 4) ☐ Interview Summary (PTO-413)
Paper No(s)/Mail Date _____
- 5) ☐ Notice of Informal Patent Application (PTO-152)
- 6) ☐ Other: _____

DETAILED ACTION

Continued Examination Under 37 CFR 1.114

1. A request for continued examination under 37 CFR 1.114, including the fee set forth in 37 CFR 1.17(e), was filed in this application after final rejection. Since this application is eligible for continued examination under 37 CFR 1.114, and the fee set forth in 37 CFR 1.17(e) has been timely paid, the finality of the previous Office action has been withdrawn pursuant to 37 CFR 1.114. Applicant's submission filed on June 1, 2006 has been entered.

Response to Amendment

2. In response to the Office Action mailed March 24, 2006, Applicants have submitted an Amendment, filed June 1, 2006, amending claims 1-2, 8-12 and 18-19, without adding new matter, and arguing to traverse claim rejections.

Response to Arguments

3. Applicant's arguments with respect to claims 1-19 have been considered but are moot in view of the new ground(s) of rejection, below.

Claim Rejections - 35 USC § 102

4. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless –

(b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

5. Claims 1-6 and 11-16 are rejected under 35 U.S.C. 102(b) as being anticipated by Hsu et al. (“Hsu”), EP 1 072 986 (published Jan 31, 2001).

Regarding claims 1 and 11, Hsu teaches a context-aware tokenizer comprising: at least one context automation module that generates a context record (contextual rules) associated with text strings of an input data stream (Fig. 15(a), 15(c) and par [77 through 82]); and

a tokenizing automation module having a token automaton that partitions said input data stream into substrings corresponding to tokens by taking context of a token into account and using it as a precondition to its recognition (see par. [42 through 45]; par. [45] teaches “The term ‘contextual rule’ refers to comparing the *context of a token* with a set of predetermined token patterns to see if there is a match. The context of a token includes the token under consideration and possibly the tokens before and/or after it”).

Regarding claims 2 and 12, Hsu teaches wherein said context automaton module comprises a left context automaton that populates (generates) said context record based on identified patterns that precede a given text string and a right context automaton that populates (generates) said context record (contextual rules) based on identified patterns that follow said given text string (Fig. 15 (a), Fig. 15(c) and par. [77 through 82]).

Regarding claims 3 and 13, Hsu teaches wherein said tokenizing automaton module maintains a data store of predefined token classes (token type) (Fig. 4 and par. [0053]), and assigns each token identified to at least one of said predefined token classes (par. [49 through 51]).

Regarding claims 4 and 14, Hsu teaches wherein said tokenizer reports information indicative of the position, length and class membership of tokens identified (The reference teaches that Fig. 5 is the text sequence segmented into tokens using the token types listed in Fig. 4, Fig. 5 and par. [54 through 55]).

Regarding claims 5 and 15, Hsu teaches wherein said tokenizing automaton defines a failure state (incorrect matches), and wherein said tokenizing automaton module monitors the occurrence of said failure state to maintain a record of the longest match (longest match corresponds to pattern results for the largest number value for $(p-n)/(p+n)$) found involving said failure state to detect a default token (broader token class) in the absence of any matching patterns taken from said context automaton module and said automation module (Fig. 17 (a), Fig. 18, element 1810 and par. [83 through 87]).

Regarding claims 6 and 16, Hsu teaches wherein said context automaton scans (reads) said input data stream (text sequence) in a left-to-right direction (a first direction) to acquire left context information and in a right-to-left direction (a second direction) to acquire right context information (par. [44 through 46]).

Claim Rejections - 35 USC § 103

6. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

Art Unit: 2626

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

7. Claims 7 and 17 are rejected under 35 U.S.C. 103(a) as being unpatentable over Hsu in view of Reps (“‘Maximal-Munch’ Tokenization in Linear Time,” ACM 1998).

Hsu fails to teach wherein said context automaton and said tokenizing automaton collectively obey a linear time operating constraint. However, Reps does teach of a context automaton and tokenizing automaton that collectively obey a linear time operating constraint (p. 263 and 267).

Therefore, it would have been obvious for one of ordinary skill in the art at the time of the applicant’s invention to supplement Hsu’s tokenizer with Reps linear time operating constraint to allow for reduction of storage utilization, as taught by Reps (p. 267).

8. Claims 8 and 18 are rejected under 35 U.S.C. 103(a) as being unpatentable over Hsu in view of Pereira et al. (Pereira), US Patent No. 5,781,884).

Hsu teaches wherein said tokenizing automaton module partitions said text strings to include token class membership information (Fig. 5 teaches the text sequence segmented into tokens using the token types listed in Fig. 4; Fig. 5 and par. [54 through 55]). Hsu lacks disclosing a text-to-speech synthesizer wherein the information from the partition influences the pronunciation of the text strings. However, Pereira teaches a text-to-speech synthesizer (TTS system) wherein information from said partitioned text strings influences the pronunciation of said text strings (col. 4, line 10 through col. 5, line 4 and col. 6, lines 20-35).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of applicant's invention to supplement Hsu's tokenizer with Pereira's text-to-speech synthesizer to allow for a multilingual system that is capable of handling a wide range of languages including Chinese or Japanese, as taught by Pereira (col. 1, lines 20-23).

9. Claims 9-10 and 19 are rejected under 35 U.S.C. 103(a) as being unpatentable over Hsu in view of Corston-Oliver et al. ("Corston-Oliver"), US Patent App. Pub. No. 2002/0138248.

Regarding claims 9-10, Hsu fails to teach a text processor coupled to a tokenizing automaton. However, Corston-Oliver teaches a tokenizing automaton (message parser; Fig. 2, element 204) coupled to said text processor (linguistic analyzer; Fig. 2, element 206) wherein input data stream (message) comprises text that lacks word unit separation symbols (Japanese). (It is well known that Japanese text does not contain word space indicators as is found in European or romance languages). Corston-Oliver also teaches said text processor operating upon said text to identify and label multi-word phrases/units for single unit treatment (Fig. 4, element 224 and par. [55 through 88]).

Therefore, it would have been obvious for one of ordinary skill in the art at the time of applicant's invention to supplement Hsu's tokenizer with Corston-Oliver's text processor to allow for text to be compressed and more easily displayed on small screens in a linguistically intelligent manner, as taught by Corston-Oliver (par. [1]).

Regarding claim 19, Hsu fails to teach generating tokenization information about said input stream (message) that includes class membership (meaning, part-of-speech) of said tokens

Art Unit: 2626

(pronoun, verb, etc.) and supplying said tokenization information to a text processor. However, Corston-Oliver teaches generating tokenization information about input stream that includes class membership of predefined tokens and supplying tokenization information to a text processor (linguistic analyzer) (Fig. 2, element 206; Fig. 4, element 222, element 224 and par. [25-27 and 35-45]).

Therefore, it would have been obvious for one of ordinary skill in the art at the time of applicant's invention to supplement Hsu's method for tokenizing with Corston-Oliver's method for supplying tokenization information to a text processor to allow for text to be compressed and more easily displayed on small screens in a linguistically intelligent manner, as taught by Corston-Oliver (par. [1]).

Conclusion

10. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure.

- US 6,327,561 to Smith *et al.* teaches customized tokenization of domain (context) specific text via rules corresponding to a speech recognition vocabulary.

11. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Eunice Ng whose telephone number is 571-272-2854. The examiner can normally be reached on Monday through Friday, 8:30 a.m. - 5:00 p.m.

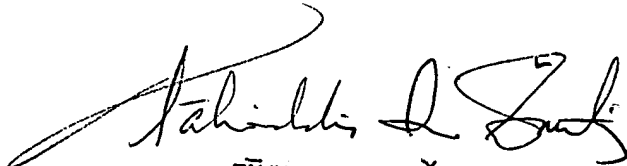
If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, David Hudspeth can be reached on 571-272-7843. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Art Unit: 2626

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

EN

July 28, 2006



TĀLVALDIS IVARS ŠMITS
PRIMARY EXAMINER